

Production of Coronavirus Nonstructural Proteins in Soluble Form for Crystallization

Yvonne Piotrowski, Rajesh Ponnusamy, Stephanie Glaser,
Anniken Daabach, Ralf Moll, and Rolf Hilgenfeld

Abstract

For biophysical investigations on viral proteins, in particular for structure determination by X-ray crystallography, relatively large quantities of purified protein are necessary. However, expression of cDNAs coding for viral proteins in prokaryotic or eukaryotic systems is often not straightforward, and frequently the amount and/or the solubility of the protein obtained are not sufficient. Here, we describe a number of protocols for production of nonstructural proteins of coronaviruses that have proven to be efficient in increasing expression yields or solubilities.

Key words: nonstructural proteins; human coronavirus; expression; solubility; crystallization.

1. Introduction

The number of viral outbreaks has increased dramatically in recent years. Since 1996, the world has seen at least one major outbreak of a new virus or of a new variant of a known virus each year. Vaccines are hardly capable of stopping such outbreaks, because even with advances in their development, they will not be available for 6 to 12 months following the first appearance of a new pathogen. Immediate containment of viral outbreaks will therefore depend on quarantine and antiviral drugs. Unfortunately, no drug treatment is available for most known

viral diseases of humans, let alone for newly emerging ones. We believe that in view of this situation, it is necessary to develop lead compounds with activity against all major families of viruses, both those that infect humans and those that so far have been restricted to animals but may cross the species barrier by zoonotic transmission. Ultimately, we should aim at discovering antiviral compounds that exhibit a relatively broad activity against a range of new viruses should they emerge.

In order to discover such leads for antiviral compounds, a sophisticated approach is necessary. Random screening for antivirals may have its merits, but the case of HIV has demonstrated that this approach has not led to a single marketed drug in the past 20 years, whereas no fewer than 26 antiretroviral drugs that have been rationally designed are on the market or in late development phases. Therefore, we follow a structure-based approach to discover new compounds with activity against RNA viruses. Methods applied include *de-novo* design, virtual screening, and structure-guided medicinal chemistry, all of which require a detailed knowledge of the structure and function of the viral target enzymes.

Coronaviruses such as human coronavirus 229E (HCoV 229E) and HCoV OC43 are responsible for variants of the common cold, while SARS-CoV caused the outbreak of severe acute respiratory syndrome in 2003, which killed more than 800 people (1). The human coronavirus NL63, first described in March 2004, has been identified as the causative agent of respiratory disease in very young children and in immunocompromised adults (2). Furthermore, HCoV-NL63 infection has been associated with laryngotracheitis (croup) (3) and Kawasaki disease, although the latter has been contested by a recent study (4).

The emergence of coronaviruses during the last decade and the wide variety of the diseases they cause—from the relatively harmless common cold to potentially lethal SARS—demonstrate the necessity of understanding the replication and transcription of these viruses, in order to find drugs that will interfere with these processes.

The 16 nonstructural proteins (Nsp) of the coronaviruses are required for genomic RNA synthesis (replication) and subgenomic RNA synthesis (transcription). They are encoded by the replicase gene, which comprises two open reading frames, ORF1a and ORF1b. ORF1a encodes the replicative polyprotein 1a (pp1a). The larger pp1ab results from a (−1) ribosomal frameshift during translation, which occurs just upstream of the ORF1a stop codon. Thus, pp1ab is the translation product of ORF1a with a large extension resulting from the ORF1b-encoded part. The polyproteins are processed into individual polypeptides by the virus-encoded proteases, one or two papain-like cysteine proteases (domains of Nsp3) and the main protease (Nsp5). Even though the function of a few nonstructural proteins is known, e.g., Nsp5 being the main protease, Nsp13

being a helicase and Nsp12 being an RNA-dependent RNA polymerase (1,5), the process of replication and transcription has not yet been elucidated. It is expected that this problem will be tackled by structural and functional studies of individual nonstructural proteins and of the complexes they form with one another and with nucleic acids.

X-ray crystallography and nuclear magnetic resonance spectroscopy are the most important methods for determination of the three-dimensional structures of proteins at high resolution. Today, more than 85% of protein structures are elucidated by X-ray crystallography. For crystallization, rather elevated concentrations of the protein under study are needed (5–20 mg/ml) and the sample should be as pure as possible (>95%). To achieve this, an efficient expression system and a reliable purification protocol have to be established. Here we present a few protocols for expression in *Escherichia coli* of genes coding for nonstructural proteins of human coronaviruses NL63 and 229E and for increasing the solubility of the recombinant protein. We also describe the purification of proteins carrying hexahistidine or GST (glutathione-S transferase) affinity tags. Furthermore, we present a preliminary crystallization experiment that has been shown to give initial hints at crystallization conditions.

2. Materials

2.1. Cultivation

1. The culture medium is prepared by dissolving 15.5 g of Invitrogen's 2× tryptone-yeast (TY) powder (1.6% tryptone, 1% yeast extract, 0.5% NaCl) in 1 liter distilled water followed by autoclaving.
2. To prepare a stock solution of ampicillin, 1 g of ampicillin (Na-salt) is dissolved in 5 ml glycerol and 5 ml distilled water and filtered through a 0.45- μ m filter. The solution is stored at -20°C .
3. The 1-M isopropyl- β ,D-thiogalactopyranoside (IPTG) stock solution is prepared in 10 ml distilled water, filtered through a 0.45- μ m filter, and stored at -20°C .

2.2. Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. a. Laemmli solution A: 30 g acrylamide, 0.3 g N,N' -bismethylenacrylamide, filled up to 100 ml with distilled water, filtered. Store at 4°C .
b. Laemmli solution B: 18.15 g Tris-HCl, filled up with distilled water to 100 ml, pH 8.8. Store at 4°C .
c. Laemmli solution C: 10 g SDS dissolved in distilled water in a total volume of 100 ml.

- d. Laemmli solution D: 3 g Tris-HCl, filled up with distilled water to a final volume of 50 ml, pH 6.8. Store at 4°C.
 - e. 10% (w/v) ammonium persulfate solution (APS): Dissolve 1 g APS in 10 ml distilled water and filter through a 0.45- μ m filter. Store the solution at 4°C.
 - f. N,N,N',N'-tetramethylethylenediamine (TEMED). Store at 4°C.
 - g. 0.1% (w/v) SDS solution: 100 mg SDS dissolved in distilled water in a total volume of 100 ml.
2. Laemmli sample buffer: 1.5 ml Laemmli solution D, 2 ml Laemmli solution C, 2 ml glycerol 87% (w/v), 3.1 ml distilled water, 0.2 ml of a 1% (w/v) bromophenol blue solution (in ethanol), freshly add one-tenth of the β -mercaptoethanol to the mixture.
 3. Mini gel running buffer: 1 g SDS, 3 g Tris, 14.4 g glycine, filled up to 1 liter with distilled water.
 4. Coomassie staining solution: 0.01% (w/v) Brilliant Blue R 250 (Sigma), 50% (v/v), methanol, 10% (v/v) acetic acid.
 5. Destaining solution: 40% (v/v) methanol, 10% (v/v) acetic acid.

2.3. Solubility Screen [after (6)]

1. Washing buffer: 10 mM Tris-HCl, pH 8.5, 100 mM NaCl, 1 mM EDTA (*see Note 1*).
2. Lysozyme solution: Dissolve 15 mg lysozyme in 1 ml of distilled water.
3. Solubility screen buffers for His-tagged protein (*see Note 2*):
 - a. 100 mM Tris, 10% (v/v) glycerol, pH 7.6
 - b. 100 mM Tris, 50 mM LiCl, pH 7.6
 - c. 100 mM HEPES, 50 mM $(\text{NH}_4)_2\text{SO}_4$, 10% (v/v) glycerol, pH 7.0
 - d. 100 mM HEPES, 100 mM KCl, pH 7.0
 - e. 100 mM Tris, 50 mM NaCl, 10% (v/v) isopropanol, pH 8.2
 - f. 100 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, 50 mM $(\text{NH}_4)_2\text{SO}_4$, 1% (v/v) Triton X-100, pH 6.0
 - g. 100 mM triethanolamine, 100 mM KCl, 10 mM DTT, pH 8.5
 - h. 100 mM Tris, 100 mM sodium glutamate, 10 mM DTT, pH 8.2
 - i. 250 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, 0.1% (w/v) CHAPS, pH 6.0
 - j. 100 mM triethanolamine, 50 mM LiCl, 5 mM EDTA, pH 8.5
 - k. 100 mM sodium acetate, 100 mM glutamine, 10 mM DTT, pH 5.5
 - l. 100 mM sodium acetate, 100 mM KCl, 0.1% (v/v) n-octyl- β -D-glucoside, pH 5.5
 - m. 100 mM HEPES, 1 M MgSO_4 , pH 7.0
 - n. 100 mM HEPES, 50 mM LiCl, 0.1% (w/v) CHAPS, pH 7.0
 - o. 100 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, 2,5 mM ZnCl_2 , pH 4.3
 - p. 100 mM Tris, 50 mM NaCl, 5 mM calcium acetate, pH 7.6
 - q. 100 mM triethanolamine, 50 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM MgSO_4 , pH 8.5
 - r. 100 mM Tris, 100 mM KCl, 2 mM EDTA, 1% (v/v) Triton X-100, pH 8.2
 - s. 100 mM sodium acetate, 1 M MgSO_4 , pH 5.5

- t. 100 mM Tris, 2 M NaCl, 0.1% (v/v) n-octyl- β -D-glucoside, pH 7.6
 - u. 100 mM Tris, 1 M $(\text{NH}_4)_2\text{SO}_4$, 10 mM DTT, pH 8.2
 - v. 100 mM sodium acetate, 50 mM LiCl, 5 mM calcium acetate, pH 5.5
 - w. 100 mM HEPES, 100 mM sodium glutamate, 5 mM DTT, pH 7.0
 - x. 100 mM triethanolamine, 100 mM sodium glutamate, 0.02% (v/v) n-octyl- β -D-glucoside, 10% (v/v) glycerol, pH 8.5
 - y. 100 mM Tris, 50 mM NaCl, 100 mM urea, pH 8.2
 - z. 100 mM triethanolamine, 100 mM KCl, 0.05% (w/v) dextran sulfate, pH 8.5
 - aa. 100 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, 50 mM $(\text{NH}_4)_2\text{SO}_4$, 0.05% (w/v) dextran sulfate, pH 6.0
 - bb. 100 mM HEPES, 50 mM LiCl, 0.1% (w/v) deoxycholate, pH 7.0
 - cc. 100 mM Tris, 100 mM KCl, 0.1% (w/v) deoxycholate, 25% (v/v) glycerol, pH 7.6
 - dd. 100 mM potassium acetate, 50 mM NaCl, 0.05% (w/v) dextran sulfate, 0.1% (w/v) CHAPS, pH 5.5
4. Solubility screen buffers for glutathione-S-transferase (GST) fusion protein (*see Note 3*): Basic buffer is PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 , 1.8 mM KH_2PO_4 , adjusted to pH 7.3 with HCl) containing:
- a. 1% Tween 20.
 - b. 1% Triton X-100.
 - c. 2 mM EDTA.
 - d. 10% glycerol.
 - e. negative control.
 - f. 5 mM DTT.
 - g. 0.1% n-octyl- β -D-glucopyranoside 0.1%.
 - h. 0.17 mM dodecylmaltoside.

2.4. Large-Scale Purification

2.4.1. Purification of the His-Tagged Protein

1. Binding buffer: Solubility screen buffer + 20 mM imidazole (*see Note 4*).
2. Elution buffer: Solubility screen buffer + 500 mM imidazole (*see Note 4*).
3. Hen egg-white lysozyme, lyophilized.
4. Complete protease inhibitor cocktail tablet (EDTA-free) (Roche).
5. HisTrap FF 1-ml column (Amersham Biosciences).
6. Amicon ultrafiltration cell.

2.4.2. Purification of the GST-Fusion Protein

1. $1 \times$ PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 , 1.8 mM KH_2PO_4 , pH 7.3) + additive (*see Note 5*).
2. Cleavage buffer: 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.5.

3. Elution buffer: 50 mM Tris-HCl, 20 mM reduced glutathione, pH 8.0.
4. Hen egg-white lysozyme, lyophilized.
5. Complete protease inhibitor cocktail tablet (EDTA-free) (Roche).
6. PreScission protease: 10 mg/ml.
7. GSTrap FF 5-ml column (Amersham Biosciences).
8. Amicon ultrafiltration cell.

2.5. Preliminary Crystallization Experiment

1. A pregreased 48-well plate from Hampton Research (VDX48 plate with sealant, cat.-no. HR3-275) and 12-mm-diameter circle cover slides.
2. Buffers: 1 M citric acid pH 4.0, 1 M citric acid pH 5.0, 1 M MES pH 6.0, 1 M HEPES pH 7.0, 1 M Tris pH 8.0, 1 M bicine pH 9.0, adjusted to the desired pH with HCl or NaOH (*see Note 6*).
3. Precipitant solutions: 4 M $(\text{NH}_4)_2\text{SO}_4$, 50% polyethylene glycol (PEG) 6000.

3. Methods

The more that is known about the target protein prior to the onset of experiments, the better. Homologous sequence alignments and secondary-structure prediction programs are helpful tools from the outset. They can provide a hint as to whether to clone the whole gene into the expression vector or to truncate it because of potentially unfolded regions, or to even go for larger parts of polyproteins rather than individual proteins. An example is shown in **Fig. 1** (*see Note 7*).

In this chapter, we present a guide to the production of sufficient amounts of soluble protein and describe it for some nonstructural proteins of HCoV NL63 and HCoV 229E. We do not describe cloning procedures, as they are common knowledge in almost every laboratory working in the field. However, we do provide a recipe for a preliminary crystallization experiment.

We start with various expression constructs of the target and the search for suitable expression conditions. All expression experiments are performed in *E. coli* strains. The cultivation experiments are done on a small scale (Section 3.3, step 2) in order to test as many parameters as possible. We show that the expression yield can differ using different temperatures, times of induction, and amounts of inductor.

Having found the conditions for optimal expression, we cultivate on a larger scale (Section 3.3, step 3), because more cells are needed for the subsequent solubility screen. The solubility screen (Section 3.4.) demonstrates that the choice of the purification tag that was made during the cloning procedure at the very beginning, as well as the use of additives, can have a huge effect on the solubility of the target protein. Furthermore, co-purification of two or more individual

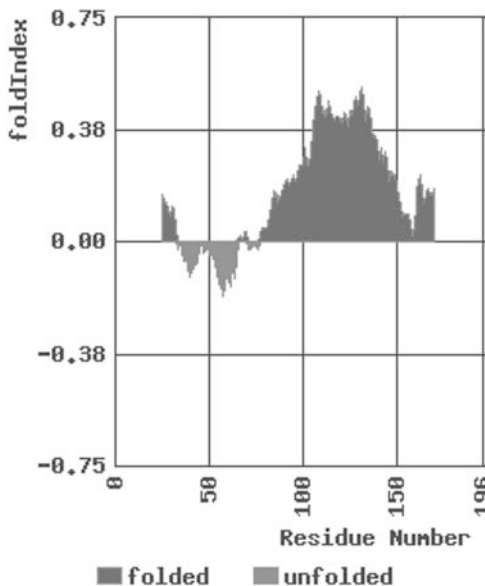


Fig. 1. FoldIndex[®] (7) prediction for HCoV-229E Nsp8. Peaks below 0.00 on the Y-axis show regions predicted to be unfolded; peaks above 0.00 indicate folded regions.

proteins can improve the solubility dramatically and might indicate that these proteins function as a complex, which we demonstrate with some figures.

Finally, we demonstrate that investing to obtain well-soluble protein is worth the effort. We perform a preliminary crystallization screen (Section 3.6.) using ammonium sulfate and PEG. Even this initial screen can sometimes yield crystals if the protein is well purified and reasonably soluble. Such crystals can usually be optimized and used in structure determination by X-ray crystallography.

3.1. SDS-PAGE

1. The SDS-PAGE instruction presented here uses the Hoefer SE250 gel system, but it can be easily modified to suit any other system.
2. Prepare a 0.75-mm-thick 15% gel by pipetting 2.78 ml Laemmli solution A, 1.38 ml Laemmli solution B, 55 μ l Laemmli solution C, 1.28 ml water, 55 μ l 10% APS, and 3.3 μ l TEMED. Pour the gel and leave some space for the stacking gel. To avoid evaporation, carefully pour 0.1% (w/v) SDS on top of the nonpolymerized gel solution. Wait 20–30 min, until the gel is polymerized.
3. Decant the 0.1% (w/v) SDS solution.
4. For the stacking gel, pipette 1 ml Laemmli solution A, 50 μ l Laemmli solution C, 1.25 ml Laemmli solution D, 2.65 ml water, 50 μ l APS, and 5 μ l TEMED. Pour the stacking gel on top of the separating gel. Insert a comb with the necessary number

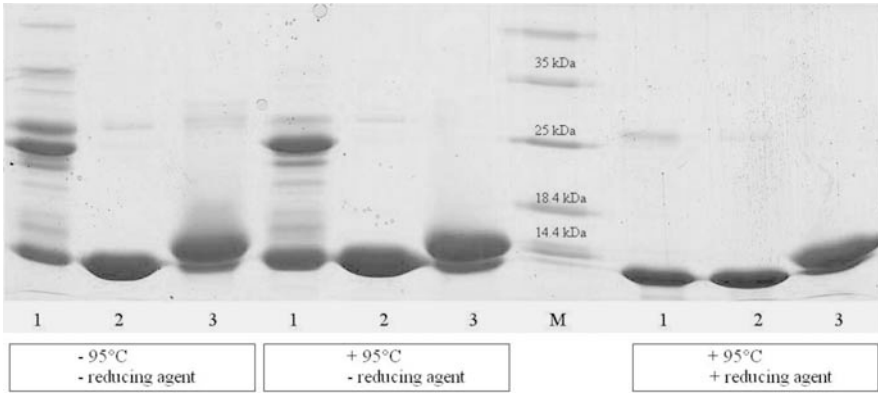


Fig. 2. SDS-PAGE of Nsp9 (14 kDa) of HCoV 229E (wildtype [containing one disulfide bridge in the dimer (1)]; mutant [Cys mutated to Ala] (2)], and SARS-CoV (3) under different conditions; reducing agent = β -mercaptoethanol; M = protein molecular-weight marker (Fermentas).

of pockets into the stacking gel. Wait 20–30 min, until the gel is polymerized. Carefully remove the comb after polymerization (*see Note 8*).

5. Attach the gel sandwich to the core that is fixed to the lower chamber in the case of the Hoefer SE250 gel system. Fill the sample wells and every upper and lower buffer chamber with running buffer. Make sure that the lower electrode is completely submerged.
6. For sample preparation, mix an equal volume of protein solution with the Laemmli sample buffer according to the size of the wells in the stacking gel. Heat the samples for 5 min at 95°C and centrifuge for 1 min at $\sim 16,000\times g$. Load the samples and the molecular weight marker into the wells (*see Note 9; Fig. 2*).
7. The gel electrophoresis is carried out for ~ 1 h (until the blue dye front is almost running out of the gel) at 20 mA/gel.
8. Transfer the gel to a basin containing the Coomassie staining solution and incubate at room temperature for 30 min.
9. Discard the staining solution in a separate waste container and destain the gel by incubating it with destaining solution at room temperature until the background is clear.

3.2. Cell Lysis (Under Denaturing Conditions) (*see Note 10*)

1. Take a 1-ml sample from the culture and measure the optical density at 660 nm (OD_{660}).
2. Centrifuge the sample in a 1.5-ml reaction tube at $\sim 16,000\times g$, 4°C for 5 min and discard the supernatant.
3. Resuspend the pellet in 67 μ l Laemmli solution D (SDS-PAGE solution).

4. Add 8 μl of 10% SDS-solution (Laemmli solution C), resulting in a concentration of 1% SDS in the final volume.
5. Shake the tube for 10 min at 950 rpm and 95°C.
6. Centrifuge the tube at $\sim 16,000\times g$, 4°C for 10 min.
7. Transfer the supernatant into a new 1.5-ml reaction tube.

3.3. Cultivation

3.3.1. Overnight Culture (5-ml Culture Volume) (see Note 11)

1. Add 5 μl of the ampicillin stock solution (100 mg/ml) to 5 ml $1\times$ TY medium (resulting in a final concentration of 100 $\mu\text{g}/\text{ml}$ ampicillin).
2. Inoculate the 5 ml of $1\times$ TY/ampicillin medium with a single colony or a mixed colony (see **Note 12**).
3. Let the culture grow overnight at 37°C, shaking at 250 rpm.

3.3.2. Cultivation for Expression

1. Prepare a 2-ml overnight culture. Let the culture grow overnight at 37°C, shaking at 250 rpm.
2. In four parallel experiments (I, II, III, and IV), inoculate 40 ml each of fresh $1\times$ TY/ampicillin medium, using 400 μl of the overnight culture (see **Note 13**).
3. Grow the cells at 37°C, shaking at 180 rpm.
4. Remove 1 ml of culture from each cultivation (“preinduction sample”) for cell lysis under denaturing conditions—when the OD_{660} reaches 0.5 for experiments I and II and 1.0 for III and IV.
5. Induce the expression by adding 40 μl to I and II and 20 μl to III and IV of the 1-M IPTG stock solution to the remaining cultures directly after the removal of the preinduction sample, resulting in a final concentration of 1 mM and 0.5 mM IPTG, respectively. Then shift cultures I and III to 20°C and leave II and IV at 37°C, shaking at 180 rpm.
6. Expression is carried out for 4 h. After this, a 1-ml sample of each culture is taken (“postinduction sample”), the OD_{660} is determined and cell lysis is performed under denaturing conditions (see **Note 14**).
7. Analyze the samples by SDS-PAGE. Take 10 μl of each sample and adjust to an OD_{660} value of 0.5, e.g., 10 μl of the preinduction sample $\text{OD}_{660} = 0.5$ (experiments I and II), or 5 μl (filled up to 10 μl with distilled water) of the preinduction sample $\text{OD}_{660} = 1.0$ (experiments III and IV). One example of the influence of the parameters tested is shown in **Fig. 3** (see **Note 15**).

3.3.3. Cultivation for Solubility Screen

1. Prepare a 5-ml overnight culture. Let the culture grow overnight at 37°C, shaking at 250 rpm.

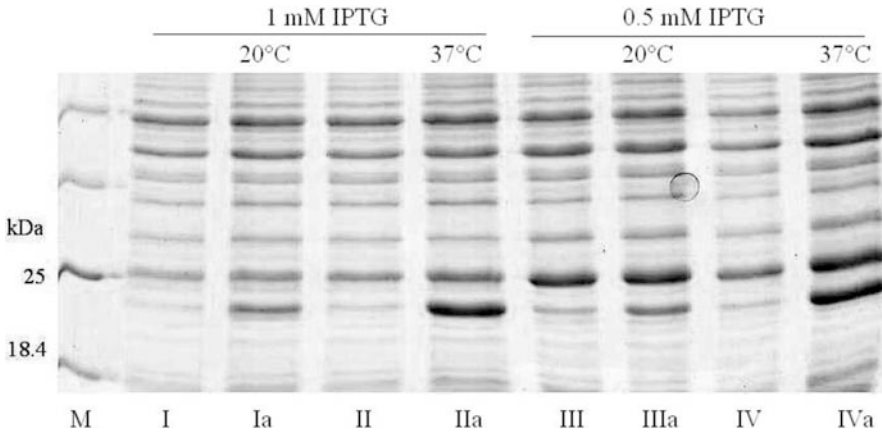


Fig. 3. Expression of the HCoV-NL63 Nsp3 X-domain with the vector pEXP1-DEST in BL21 (DE3) pLysS, containing a sequence encoding a 6xHis-tag (total molecular mass of 23 kDa). M = protein molecular-weight marker (Fermentas); I: 1 mM IPTG, 20°C, II: 1 mM IPTG, 37°C; III: 0.5 mM IPTG, 20°C, IV: 0.5 mM IPTG, 37°C; before (I–IV) and after induction (Ia–IVa).

2. Centrifuge the overnight culture at $\sim 5200 \times g$, at room temperature for 10 min. Discard the supernatant (*see Note 16*).
3. Resuspend the pellet in a few ml of fresh $1 \times$ TY/ampicillin medium.
4. Transfer the resuspended cells to 500 ml $1 \times$ TY medium, containing 500 μ l ampicillin stock solution.
5. Grow the cells at 37°C, shaking at 170 rpm, until they reach an $OD_{660} = 0.5$ (*see Note 17*).
6. Induce the expression of the desired gene by adding 500 μ l of 1-M IPTG to get a final concentration of 1 mM IPTG.
7. Let the cultures grow for 4 h at 20°C and 37°C, respectively.
8. Harvest the cells by centrifugation for 15 min in a suitable beaker at $\sim 7200 \times g$ and 4°C. Discard the supernatant, but leave about 25 ml to resuspend the pellet. Transfer the suspension to a 50-ml Sarstedt tube.
9. Centrifuge the suspension again at $\sim 5200 \times g$, 4°C for 15 min. The pellets can be stored at -20°C for later use.

3.4. Solubility Screen [Modified after (6)]

1. Resuspend and wash the pellet of the 500-ml cultivation with 30 ml of washing buffer (*see Notes 18 and 19*).
2. Divide the 30-ml suspension into 30×1.5 -ml reaction tubes, each containing 1 ml of the resuspended cells.
3. Centrifuge the tubes at $\sim 16,000 \times g$, 4°C for 10 min and discard the supernatant.

4. If the protein is produced with a His-tag, resuspend the cells in 1 ml of one of the lysis buffers listed in Section 2.3 (point 3). If the gene of interest is expressed as a GST fusion protein, resuspend the pellet in 1 ml of one of the buffers given in Section 2.3 (point 4).
5. To each tube add 10 μ l of lysozyme stock solution (resulting in a final concentration of 0.15 mg/ml lysozyme) and vortex the solution for a few seconds.
6. Lyse the cells on ice by 1-min sonication [Branson Sonifier Cell Disruptor B15 (output control 4, % duty cycle 50)].
7. After sonication, centrifuge the tubes at $\sim 16,000 \times g$, 4°C for 30 min.
8. Transfer the supernatant into a new 1.5-ml reaction tube.
9. Analyze 10 μ l of each of the 30 crude cell extracts by SDS-PAGE.
10. As controls, the complete extracts made by cell lysis under denaturing conditions before and after induction of gene expression (resulting in a final OD₆₆₀ = 0.5) are analyzed in parallel by SDS-PAGE. The influence of the purification tag on the solubility of the protein is demonstrated in **Fig. 4**. While the protein produced with the 6 \times His-tag is more or less insoluble, the protein fused to GST shows high solubility. In contrast, **Fig. 5** illustrates that instead of changing the construct, additives can also increase the amount of soluble protein (see **Note 20**).

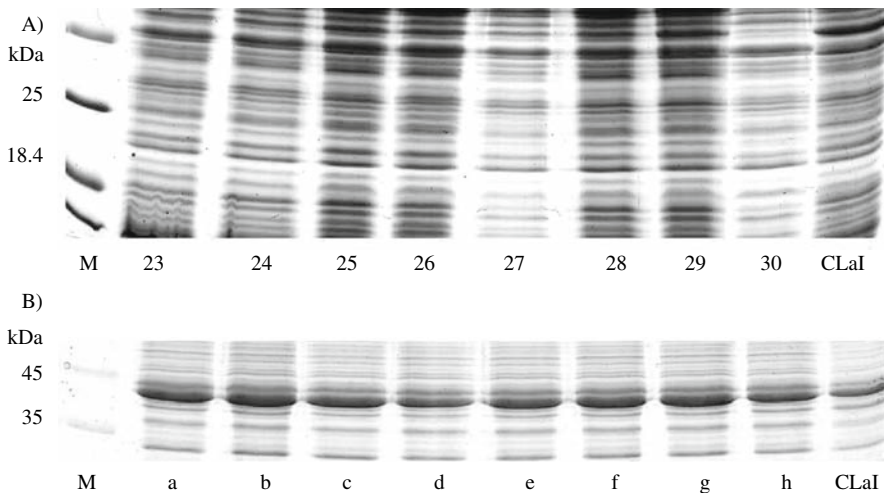


Fig. 4. Solubility screen of HCoV-NL63 Nsp3 X-domain produced with a 6xHis-tag (A, 23 kDa) and as a GST-fusion protein (B, 46 kDa). M = protein molecular-weight marker (Fermentas): (A) 23–30: buffers as listed in Section 2.3 (point 3); (B) a–h: buffers as listed in Section 2.3 (point 4); CLaI: cell lysis under denaturing conditions after induction.

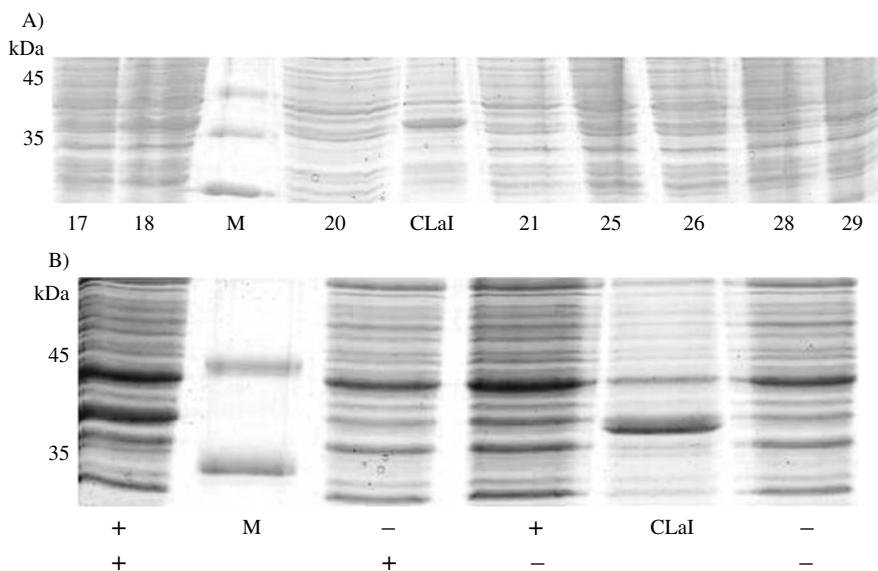


Fig. 5. Solubility screen of HCoV-NL63 Nsp1 as a GST-fusion protein (40 kDa) based on the described method for His-tagged proteins: (A) solubility screen with selected buffers (Section 2.3, point 3); (B): test for the necessity of additives based on buffer 18, upper labels: Triton X-100, lower labels: EDTA (+ in presence, – in absence); M: protein molecular-weight marker (Fermentas); CLaI: cell lysis under denaturing conditions after induction.

3.5. Large-Scale Purification (see Note 21)

3.5.1. Purification of the His-Tagged Protein

1. Resuspend the cells of a 1-liter cultivation in 25 ml binding buffer that contains 10 mg lysozyme and one-half of a tablet of complete protease inhibitor cocktail.
2. Incubate the solution for 30 min on ice.
3. Break the cells using a French press at 25,000 psi (see Note 22).
4. Perform an ultracentrifugation for 1 h at $\sim 150,000 \times g$ and 4°C .
5. Filter the supernatant through a $0.45\text{-}\mu\text{m}$ filter.
6. Apply the supernatant using a peristaltic pump with a flow of 1 ml/min on a HisTrap FF 1-ml column that has been equilibrated with the binding buffer (see Note 23).
7. Wash the column with binding buffer (1 ml/min) until the OD_{280} reaches a steady baseline.
8. Elute the protein with a flow of 1 ml/min using a step gradient of 20 mM (binding buffer) to 500 mM imidazole (elution buffer) over 50 column volumes (see Note 24). Collect the flow-through in 2-ml fractions.
9. Analyze the fractions that show an absorbance at 280 nm by SDS-PAGE.

10. Pool the fractions of highest purity and dialyze the solution overnight against an appropriate buffer at 4°C (*see Note 25*).
11. Concentrate the protein solution to 5–20 mg/ml using an Amicon ultrafiltration cell (*see Note 26*).

3.5.2. Purification of the GST-Fusion Protein

1. Resuspend the cells of a 1-liter cultivation in 25 ml 1 × PBS buffer + additive (Section 2.4.2., point 1) containing 10 mg lysozyme and one-half of a tablet of complete protease inhibitor cocktail.
2. Incubate the solution for 30 min on ice.
3. Break the cells using a French press at 25,000 psi (*see Note 22*).
4. Centrifuge for 1 h at $\sim 150,000 \times g$ and 4°C.
5. Filter the supernatant through a 0.45- μm filter.
6. Apply the supernatant using a peristaltic pump with a flow of 0.5 ml/min on a GSTrap FF 5-ml column that has been equilibrated with 1 × PBS buffer (*see Note 23*).
7. Wash with 1X PBS buffer (2 ml/min) until the OD₂₈₀ reaches a steady baseline.
8. Wash the GSTrap 5-ml column with ten column volumes (2 ml/min) cleavage buffer.
9. Prepare the PreScission protease mix by mixing 50 μl PreScission protease with 4.95 ml cleavage buffer (*see Note 27*).
10. Load the PreScission protease mix onto the column.
11. Seal the column and incubate overnight at 4°C.
12. Elute the protein with 20 ml cleavage buffer (0.5 ml/min). Collect the flow-through in 0.5-ml fractions.
13. Remove the GST protein and the PreScission protease, which is itself fused to GST, by washing the column with 2 ml/min elution buffer. Collect the flow-through in 2-ml fractions.
14. Analyze the fractions containing the protein under study, the GST, and the PreScission protease by SDS-PAGE.
15. Pool the fractions containing the protein of highest purity and concentrate to 5–20 mg/ml using an Amicon ultrafiltration cell (*see Note 26*).

3.6. Preliminary Crystallization Experiment

1. Fill the wells of the pregreased Hampton research 48-well plate (VDX48 plate with sealant) with 250 μl of the reservoir solution according to the pipetting scheme shown in **Table 1**, using a 1-M stock solution of the buffer, a 4-M stock solution of $(\text{NH}_4)_2\text{SO}_4$, and 50% PEG 6000 (*see Note 28*).
2. Centrifuge the solution of the purified protein for 10 min at $16,000 \times g$ and 4°C.
3. Pipette 1 μl of the protein solution on a circular cover slide (12 mm diameter) and add 1 μl of the reservoir solution of the corresponding well.

Table 1
Pipet Scheme for the Preliminary Crystallization Experiment with Amonium Sulfate and PEG 6000

		(NH ₄) ₂ SO ₄								PEG 6000			
		0.8 M	1.6 M	2.4 M	3.2 M	5%	10%	20%	30%				
		50 μ l	100 μ l	150 μ l	200 μ l	25 μ l	50 μ l	100 μ l	150 μ l				
		1	2	3	4	5	6	7	8				
pH 4.0	H ₂ O \nearrow	175 μ l	125 μ l	75 μ l	25 μ l	200 μ l	175 μ l	125 μ l	75 μ l				
	A												
pH 5.0	B	175 μ l	125 μ l	75 μ l	25 μ l	200 μ l	175 μ l	125 μ l	75 μ l				
	25 μ l												
pH 6.0	C	175 μ l	125 μ l	75 μ l	25 μ l	200 μ l	175 μ l	125 μ l	75 μ l				
	25 μ l												
pH 7.0	D	175 μ l	125 μ l	75 μ l	25 μ l	200 μ l	175 μ l	125 μ l	75 μ l				
	25 μ l												
pH 8.0	E	175 μ l	125 μ l	75 μ l	25 μ l	200 μ l	175 μ l	125 μ l	75 μ l				
	25 μ l												
pH 9.0	F	175 μ l	125 μ l	75 μ l	25 μ l	200 μ l	175 μ l	125 μ l	75 μ l				
	25 μ l												

4. Turn the cover slide upside down and place it on top of the well.
5. Carefully press on the cover slide to fix it onto the well.
6. Store the plate at 12°C (*see Note 29*).
7. Initially, check the plate daily, then weekly, and later monthly (*see Note 30*).

4. Notes

1. All buffers should be filtered through a 0.45- μm filter and stored at 4°C to decrease the risk of contamination.
2. The buffers for the His-tagged protein are chosen according to the literature (6). These buffers vary in composition (salts, additives) and pH stabilizing the protein and/or improving solubility.
3. According to the *GST Gene Fusion System Handbook* (Amersham Biosciences), PBS is used as a buffer for purification of the fusion protein. Additives are selected with properties of the protein in mind, i.e., isoelectric point, hydrophathy, etc.
4. The buffer that gave the best result in the solubility screen (Section 3.4) is the basic buffer. Thus, the binding buffer contains the basic buffer plus 20 mM imidazole and the elution buffer consists of the basic buffer plus 500 mM imidazole. The basic buffer resulting from the solubility screen should be compatible with the Ni^{2+} sepharose, e.g., it should not contain EDTA, as this would strip the Ni^{2+} from the column.
5. The buffer used here for cell lysis is the 1 \times PBS buffer, containing the additive that gave the best result in the solubility screen (Section 3.4).
6. The calculated amounts of the substances for the stock solutions are dissolved in distilled water and filtered through a 0.45- μm filter. For solutions that cannot be filtered because of high viscosity, e.g., the 50% PEG 6000 solution, the powder can be dissolved in prefiltered distilled water. The pH of the buffers is adjusted to the desired pH with NaOH or HCl.
7. In the case of Nsp8 of HCoV 229E, the FoldIndex[©] program (7) predicts an unfolded region at the N terminus of the protein (**Fig. 1**). Actually, this part is involved in the interaction with Nsp7, as revealed by the crystal structure of the Nsp7:Nsp8 complex of SARS-CoV (8).
8. Be sure that the gel gets polymerized; otherwise artifact bands will be observed when the protein is subjected to gel electrophoresis.
9. Depending on the protein, heating and mercaptoethanol reduction may influence the electrophoretic mobility. If the observed molecular mass differs from the one expected from the polypeptide sequence, other temperature conditions should be tried. SDS-PAGE under nonreducing, compared to reducing, conditions gives a hint concerning disulfide bonds present in the protein, as shown in **Fig. 2**.
10. During the cell lysis under denaturing conditions, the cells are totally destroyed and the proteins are unfolded. Thus, this sample represents the positive control on an SDS-PAGE gel, displaying the total amount of proteins in the cell.

11. The overnight culture is done to have an optimal growth of the cells. After the overnight incubation, the cells are adapted to the medium and strong enough to be transferred into a larger volume of medium. If the colony from the agar plate is directly transferred to a larger volume, it takes a longer time until the cells reach the desired cell density and by then the risk of contamination increases.
12. The decision as to whether to pick a single colony or a mixture of colonies depends on the stability of the clones. Using a mixture of colonies can increase the chance of finding an expression clone. In the same way, it can increase the risk of a failed expression, as the cells can contain damaged plasmids. Thus, several single colonies have to be analyzed individually. Nevertheless, checking out ten single colonies can result in no expression, whereas one or two mixtures of colonies may include good expression clones.
13. It is not absolutely necessary to centrifuge the overnight culture in order to remove degraded ampicillin and expressed β -lactamase, as the volume of the overnight culture is really small.
14. To go for a higher expression rate, additional parameters can be tested, e.g., different *E. coli* strains [BL21(DE3), Tuner pLacI (DE3), C41 (DE3)] and media (Luria Bertani, rich medium), as well as different incubation times.
15. From comparison with the intensity of the host proteins, the amount of the target protein produced can be estimated. Weak bands of the heterologously produced protein in a high background of *E. coli* proteins are difficult to detect. In any case, correct expression has to be confirmed by Western blot methods (works well for crude lysates, if the protein of interest shows up with a high yield), N-terminal sequencing, and mass spectrometry analysis (for purified recombinant protein).
16. Unlike cultivation for the expression analysis (Section 3.3.2), it is recommended to centrifuge the overnight culture to remove the degraded ampicillin and the expressed β -lactamase to avoid contamination during cultivation.
17. The expression parameters, such as induction time (OD_{660}), amount of inducer, and duration of the expression, are adjusted to the results from the initial cultivation (see Section 3.3.2). Our example is from cultures I and II, shown in **Fig. 3**.
18. Even if a better expression condition is found at 20°C, a cultivation should also be performed at 37°C, in order to test whether the protein produced in lower yields is more soluble than the one obtained in higher amounts.
19. Starting from the first step, everything should be done on ice and at 4°C to avoid protein degradation. In addition, the use of protease inhibitors, for instance, complete protease inhibitor cocktail (Roche), may be advisable.
20. Sometimes much effort can be invested in finding a buffer for cell lysis to obtain soluble protein. Depending on the function of the viral protein in replication or transcription, it may need another nonstructural protein to form a complex for proper folding. **Figure 6** demonstrates the independent purifications of Nsp7 and Nsp8 from HCoV 229E. Whereas Nsp7 is soluble, Nsp8 is not. However, by combining the pellets and purifying the two together, the solubility of Nsp8 increases dramatically as shown in **Fig. 7**. Recently, the crystal structure of a

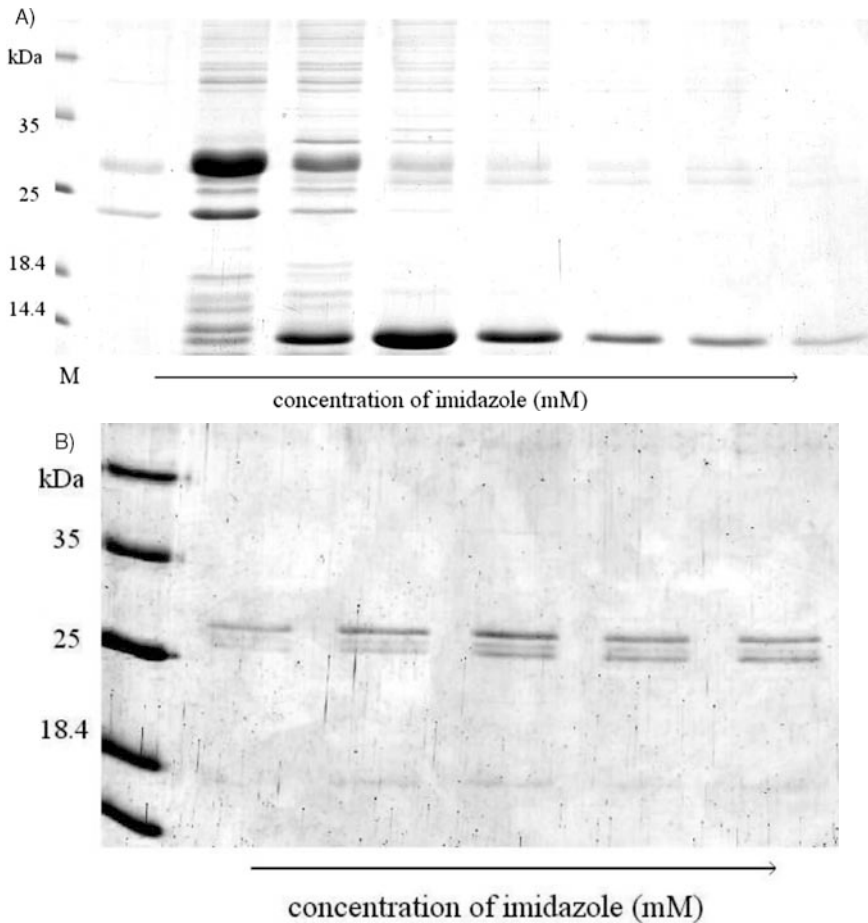


Fig. 6. Purification of His-tagged HCoV-229E Nsp7 (A) and Nsp8 (B). Elution of the proteins bound to a Ni^{2+} -NTA affinity column with increasing concentrations of imidazole. Nsp7 (11 kDa) can be eluted by small amounts of imidazole, whereas Nsp8 (22 kDa) cannot.

hexadecameric (8:8) Nsp7:Nsp8 complex from the SARS coronavirus has been determined (8).

21. Once the expression and cell-lysis conditions have been optimized to yield large amounts of soluble protein, biochemical and structural studies can be started. The protein is purified from a 1-liter cultivation (according to Section 3.3.3., with the necessary adjustments for the larger cultivation volume) using affinity chromatography: (i) *Affinity Chromatography: Principle and Methods Handbook*; (ii) *GST Gene Fusion System Handbook*: Amersham Biosciences). This

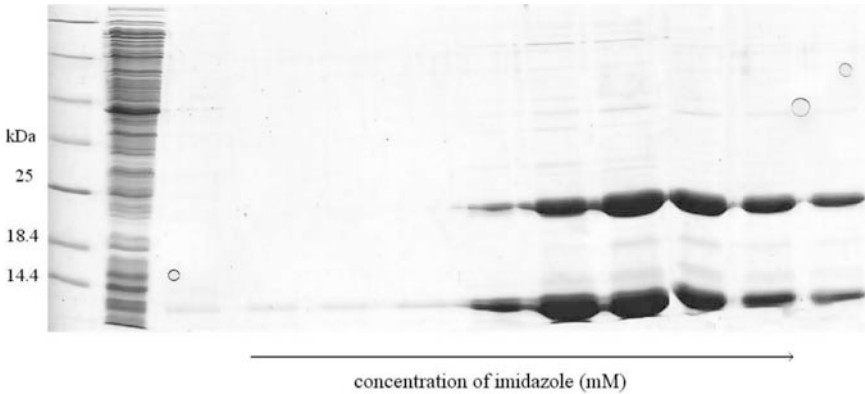


Fig. 7. Purification of HCoV-229E Nsp7 and Nsp8 in combination. The His-tagged proteins are eluted from the Ni^{2+} -NTA affinity column with increasing concentrations of imidazole. Both Nsp7 (11 kDa) and Nsp8 (22 kDa) can be eluted in larger amounts than in the single experiments (compare **Fig. 6**).

can be followed by ion-exchange or size-exclusion chromatography to remove more impurities.

22. French press is just one possibility for breaking the cells. Sonication or several thaw/freeze cycles can also be done for cell disruption.
23. Instead of a peristaltic pump, an automatic device can be used as well. For instance, the Äkta Prime (Amersham Biosciences) is easy to use for simple purification of proteins on a laboratory scale.
24. If the purification can be done with automated equipment instead of a step gradient, the protein can be eluted using a linear gradient. This gradient may separate proteins with similar binding strengths.
25. It depends on the purity of the eluted protein whether the dialysis can be done directly afterward or if a second purification step (ion-exchange or size-exclusion chromatography) is advisable to remove more impurities. Experience shows that sometimes only the purest fraction will give high-quality crystals; therefore, inclusion of too much of the flanks of the protein peak should be avoided.
26. For concentrating the protein solution, methods other than ultrafiltration can be used, depending on what is established in the laboratory.
27. In the protocol given here, the GST fusion protein contains a cleavage site for PreScission protease. Other fusion constructs may require different proteases (e.g., factor Xa) for processing. In the example described here, the protein of interest is cleaved off from the GST fusion by the PreScission protease directly on the column. It is also possible to do the cleavage in a reaction tube after elution of the GST fusion protein with the elution buffer. The fractions containing the GST fusion protein are pooled and dialyzed against the cleavage buffer before starting the cleavage. After overnight incubation, the solution is applied onto the GSTrap 5-ml column and washed with cleavage buffer. The flow-through

- contains the protein. Furthermore, cleavage of the fusion protein is not always required. A number of target-fusion crystal structures have been reported. In some cases, fusion proteins can be useful tools for crystallization (9,10).
28. In addition to the mentioned preliminary crystallization experiment with ammonium sulfate and PEG 6000 other precipitants can be tested as well, for instance, sodium chloride and 2-methyl-2,4-pentanediol (MPD) (11). There are commercial crystallization kits available. Sigma-Aldrich offers crystallization kits (basic and extension) for proteins that are based on the original screening protocol of Jancarik and Kim (12), which utilizes crystallization conditions that have worked before for various proteins and allows testing of a large range of buffers, pH, additives, and precipitants. The protocol can be used for setting up the crystallization experiment by hand, or alternatively, by pipetting robots that allow screening of a larger number of possible crystallization conditions (high-throughput experiment) with less volume of the protein solution and less effort.
 29. Temperature also has an influence on the crystallization of the protein. If there is enough protein solution available, other temperatures such as 4°C, 25°C, or 37°C should be tested.
 30. **Figure 8** shows crystals of the HCoV-NL63 Nsp3 X-domain obtained from the preliminary crystallization experiment (Section 3.6). A solution containing 1 µl protein (7 mg/ml) and 1 µl reservoir solution was equilibrated against 250 µl reservoir solution of well B7 (20% PEG 6000, 0.1 M citric acid pH 5.0). The crystal diffracted to 3.5 Å resolution. Within this screen, crystals could be obtained if PEG 6000 (10–20%) was used as precipitant in a range of pH 7.0 to 9.0, whereas (NH₄)₂SO₄ had no effect on the crystallization of the protein. The reservoir solution has been optimized and crystals have been grown that

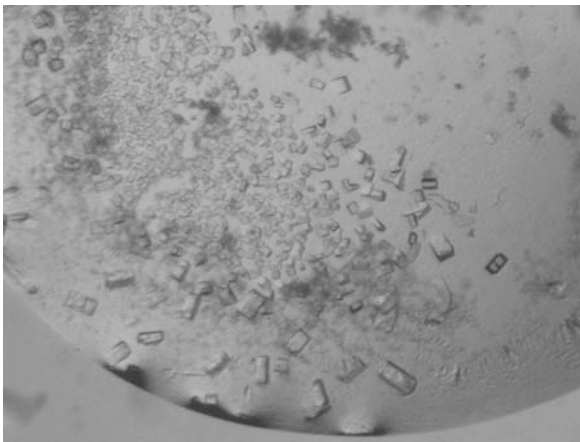


Fig. 8. Crystals of HCoV-NL63 Nsp3 X-domain. The reservoir solution contains 20% PEG 6000, 0.1 M citric acid pH 5.0.

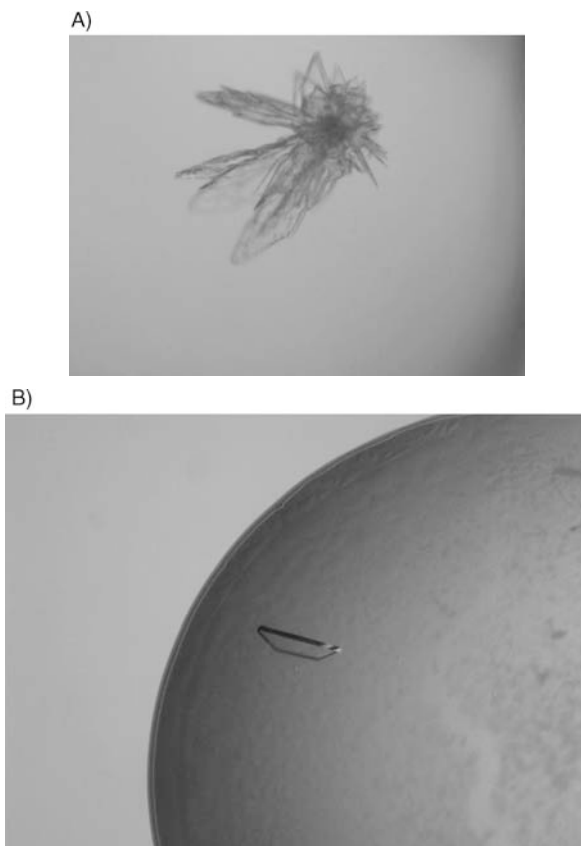


Fig. 9. Crystals of HCoV-229E Nsp9: (A) First crystallization trials yield intergrown crystals (B) Addition of 5% MPD to the reservoir buffer results in single crystals.

diffract to 1.8 Å resolution. Optimization of the crystallization conditions can also include additives such as 2-methyl-2,4-pentanediol (MPD). **Figure 9** shows intergrown crystals of Nsp9 of HCoV 229E and a single crystal grown in the presence of 5% MPD. Keep in mind, that each component of the reservoir solution can either promote or hinder the crystallization experiment. So, the choice of the buffer, salts, or organic additives should be very carefully considered and investigated (**13**).

Acknowledgments

The work described here is being supported, in part, by the VIZIER (“Comparative Structural Genomics of Viral Enzymes Involved in Replication”) project of the European Commission (contract no. LSHG-CT-2004-511960; <http://www.vizier-europe.org>), and by SEPSDA (contract no. SP22-CT-2004-003831, www.sepsda.eu). Support by the Sino-German Center on the Promotion of

Research, Beijing, is gratefully acknowledged. R. H. thanks the Fonds der Chemischen Industrie for continuous support. We thank L. van der Hoek and J. Ziebuhr for the cDNA.

References

1. Groneberg, D. A., Hilgenfeld, R., and Zabel, P. (2005) Molecular mechanisms of severe acute respiratory syndrome (SARS). *Respir. Res.* **6**, 8–23.
2. van der Hoek, L., Pyrc, K., and Berkhout, B. (2006) Human coronavirus NL63, a new respiratory virus. *FEMS Microbiol. Rev.* **30**, 760–773.
3. van der Hoek, L., Sure, K., Ihorst, G., Stang, A., Pyrc, K., Jebbink, M. F., Petersen, G., Forster, J., Berkhout, B., and Überla, K. (2005) Croup is associated with the novel coronavirus NL63. *PLoS Med.* **2**, 764–770.
4. Dominguez, S. R., Anderson, M. S., Glodé, M. P., Robinson, C. C., and Holmes, K. V. (2006) Blinded case-control study of the relationship between human coronavirus NL63 and Kawasaki syndrome. *J. Infect. Dis.* **194**, 1697–1701.
5. Stadler, K., Massignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H. D., and Rappuoli, R. (2003) SARS—Beginning to understand a new virus. *Nature Rev. Microbiol.* **1**, 209–218.
6. Lindwall, G., Chau, M.-F., Gardner, S. R., and Kohlstaedt, L. A. (2000) A sparse matrix approach to the solubilization of overexpressed proteins. *Protein Eng.* **13**, 67–71.
7. Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silmann, I., and Sussman, J. L. (2005) FoldIndex[®]: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**, 3435–3438.
8. Zhai, Y., Sun, F., Li, X., Pang, H., Xu, X., Bartlam, M., and Rao, Z. (2005) Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nature Struct. Mol. Biol.* **12**, 980–986.
9. Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P., and Kobe, B. (2003) Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **12**, 1313–1322.
10. Hogg, T., and Hilgenfeld, R. (2007) Protein crystallography in drug design. In: Tayler, J. B., and Triggle, D. J. (eds.) *Comprehensive Medicinal Chemistry II*. Elsevier, Amsterdam, pp. 875–900.
11. McPherson, A. (1999) Strategies and special approaches in growing crystals. In *Crystallization of Biological Macromolecules*. Cold Spring Harbor Laboratory Press, New York, pp. 271–329.
12. Jancarik, J., and Kim, S.-H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* **24**, 409–411.
13. Mesters, J. R. (2007) Practical protein crystallization. In: Drenth, J. (ed.) *Principles of Protein X-Ray Crystallography* 3rd Ed.. Springer, New York, pp. 297–304.